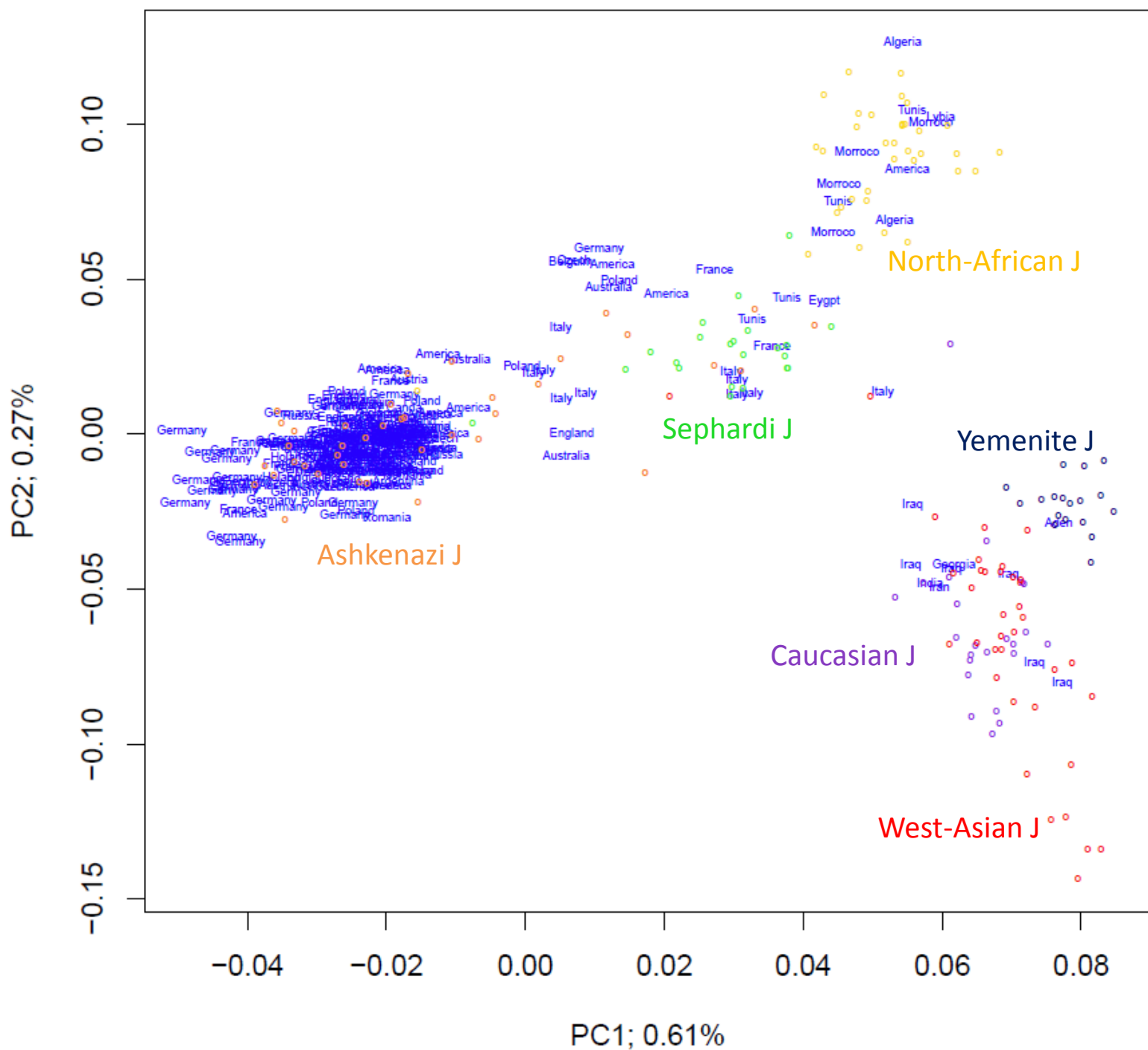
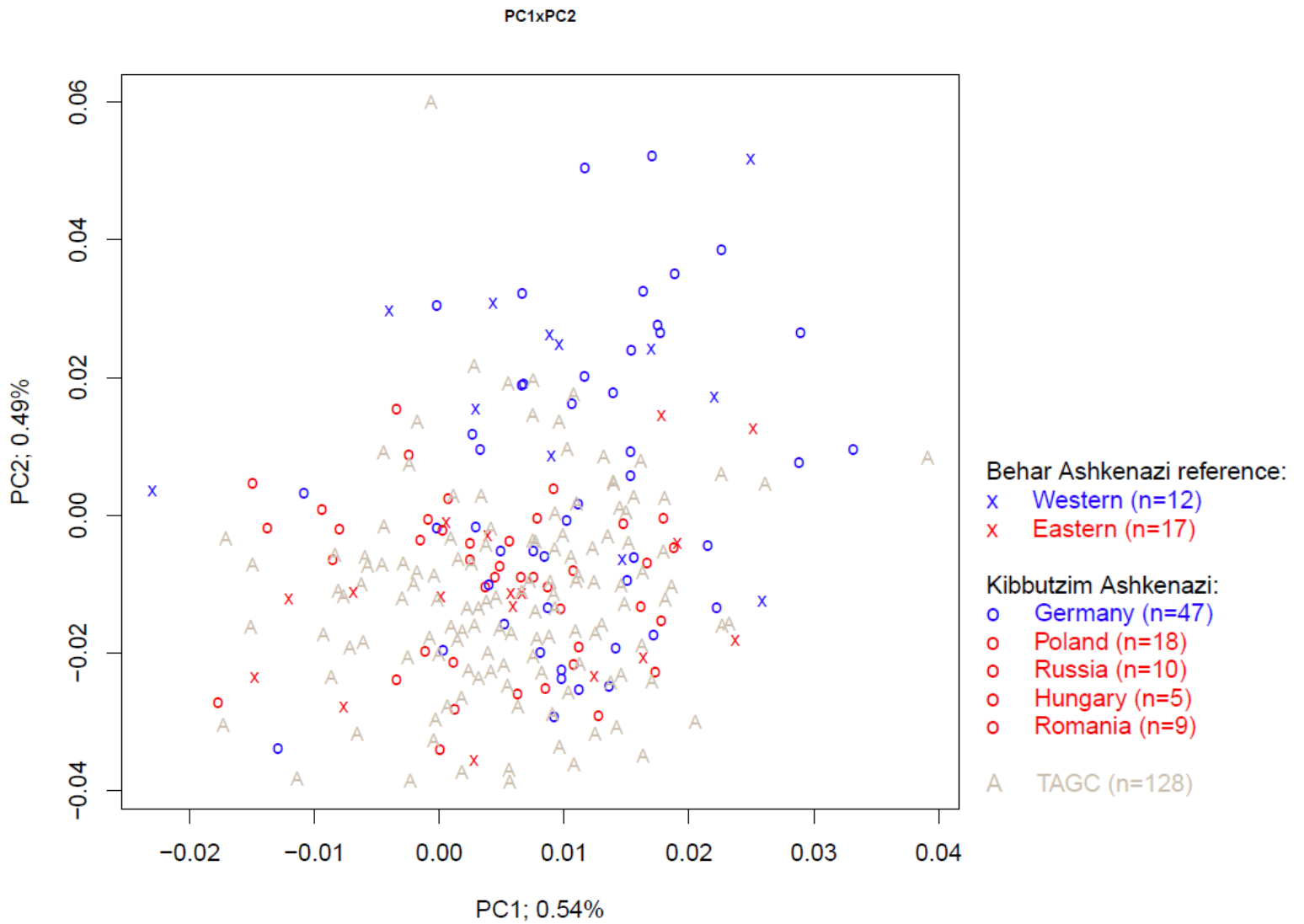


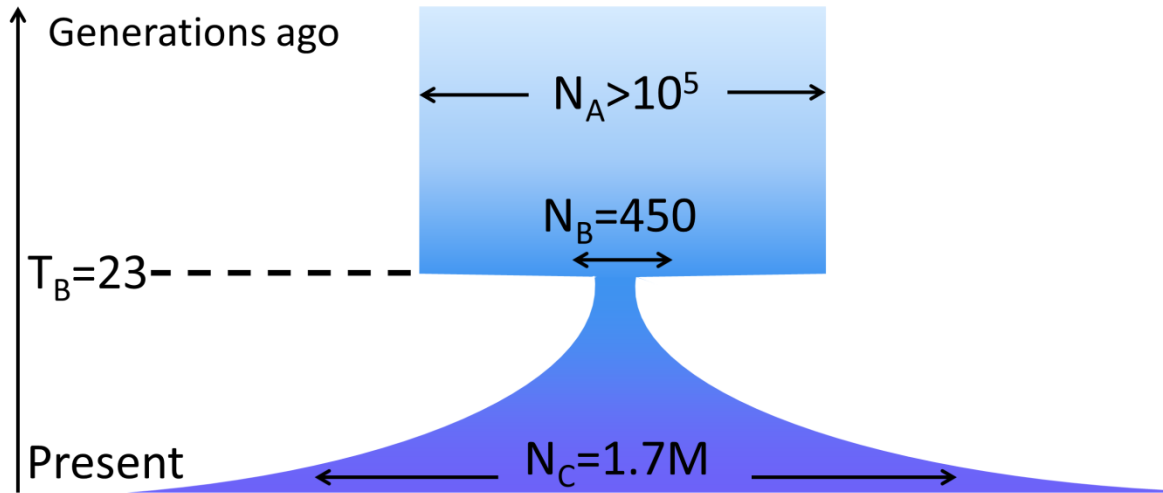
**Supplementary Figure 1:** A PCA plot for the Jewish populations. Reference samples are as in Figure 1. Samples from The Ashkenazi Genome Consortium (TAGC) are shown in beige "A"s (n=128). The KFS samples are marked as blue cross marks (n=901), demonstrating that the ancestry of the KFS samples is predominantly (though not exclusively) Ashkenazi Jewish.



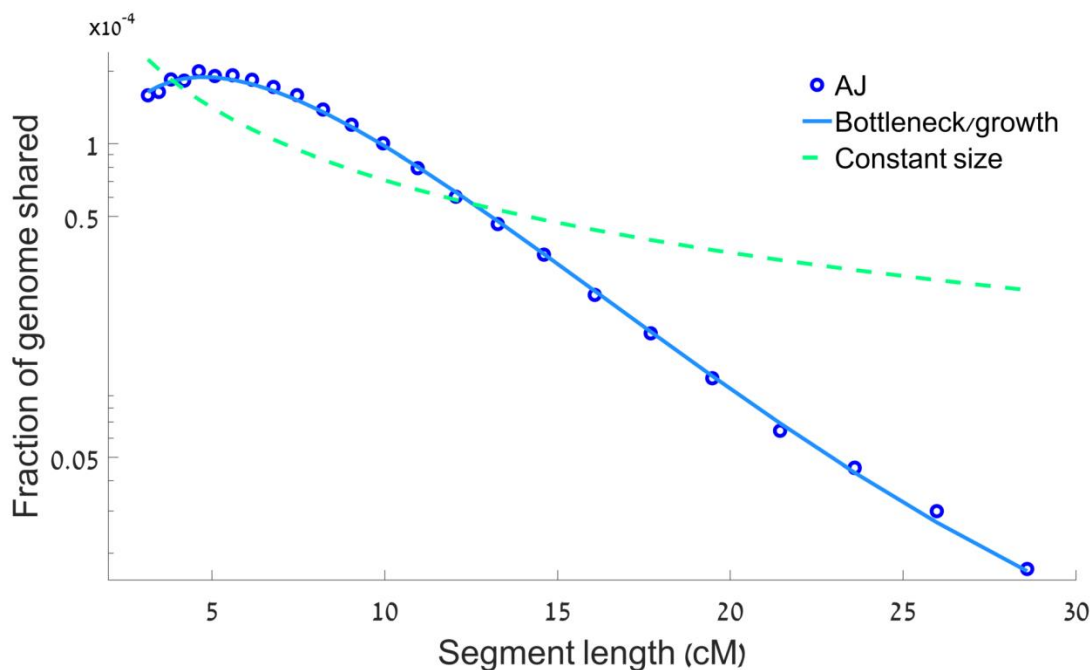
**Supplementary Figure 2:** A PCA plot of Jewish reference populations (n=174, colored circles) and the KFS (n=247). KFS samples are designated by their country of birth, in blue font. Individuals with  $PC1 < 0$ , whom we designated as having an Ashkenazi Jewish genetic ancestry, were indeed predominantly from Europe.



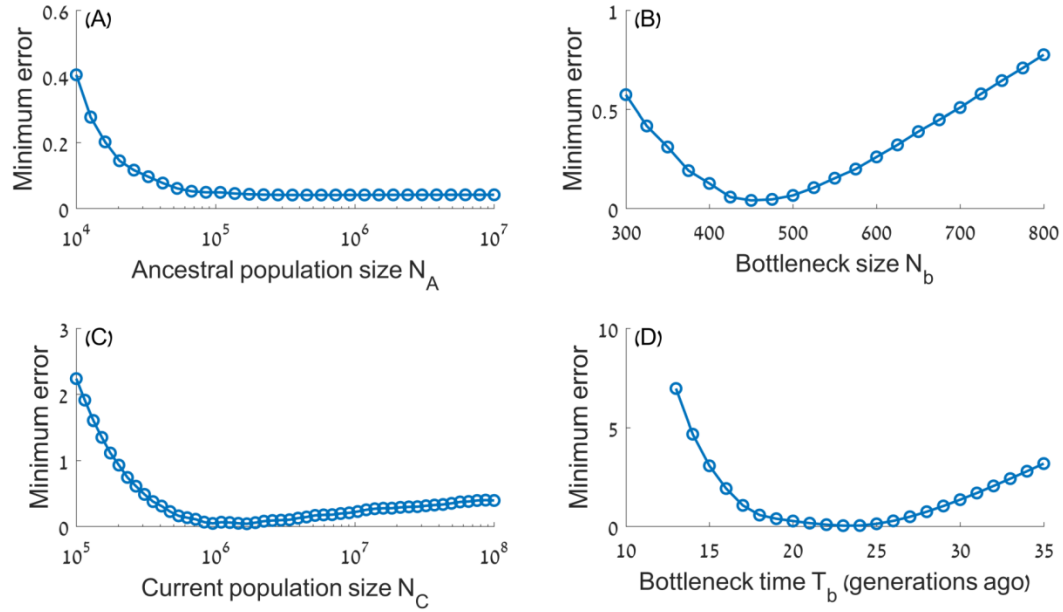
**Supplementary Figure 3:** A PCA plot of the Ashkenazi Jewish individuals. The samples of Behar et al., 2013 are designated by cross marks. Samples from the KFS are designated by circles. For both datasets, Western and Eastern AJ samples are colored blue and red, respectively. Samples from The Ashkenazi Genome Consortium (TAGC) are marked with “A”s. The plot demonstrates that Western AJ have genetic ancestry partly distinct from Eastern AJ, as many Western AJ cluster in the top-right quadrant.



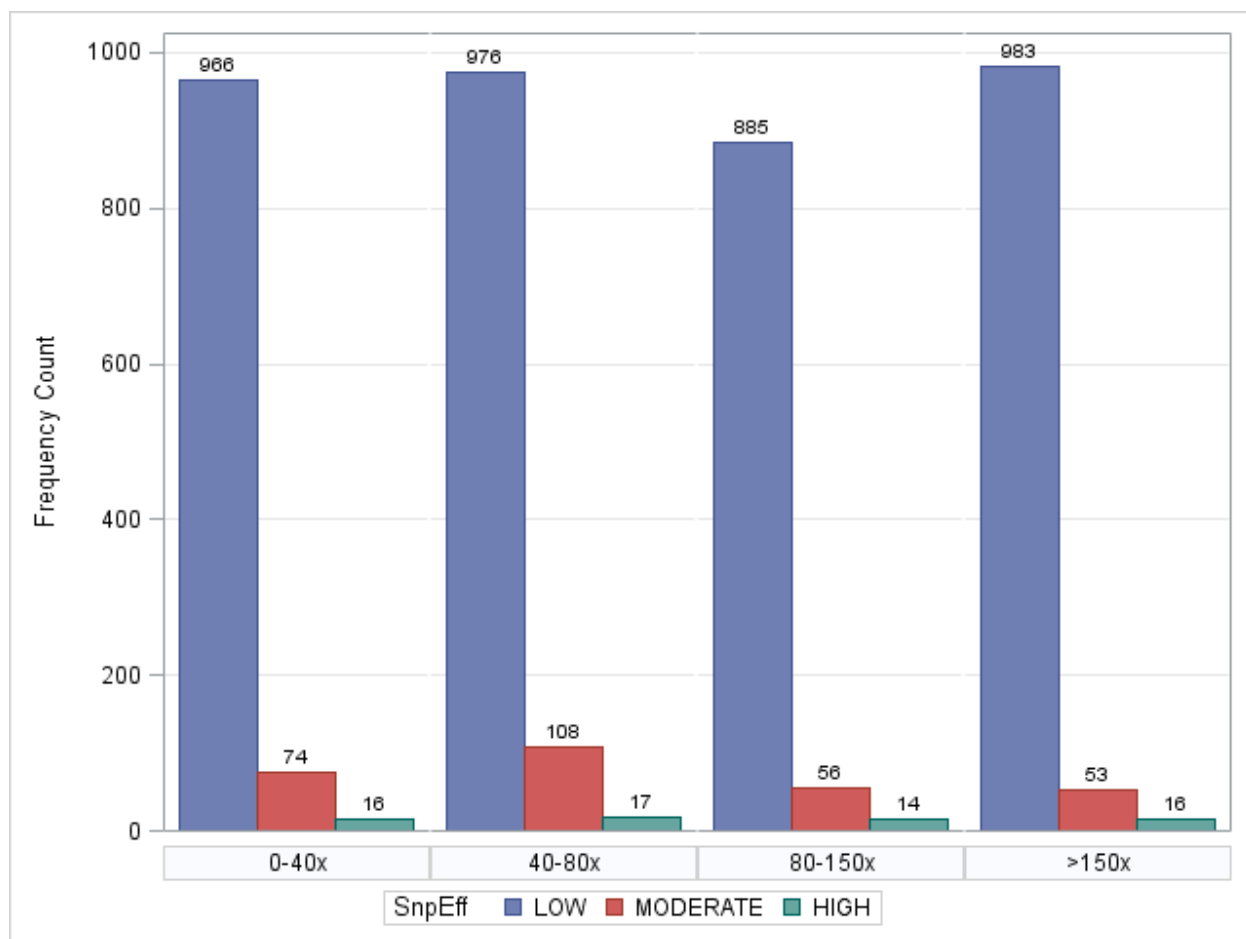
**Supplementary Figure 4:** The best fit demographic model for the Ashkenazi individuals of the Kibbutzim study. Horizontal arrows correspond to effective population sizes (number of diploid individuals). Note that the model structure (bottleneck/growth) was assumed, rather than inferred. We could not obtain a precise estimate of the ancestral population size  $N_A$  (see Supplementary Figure 6).



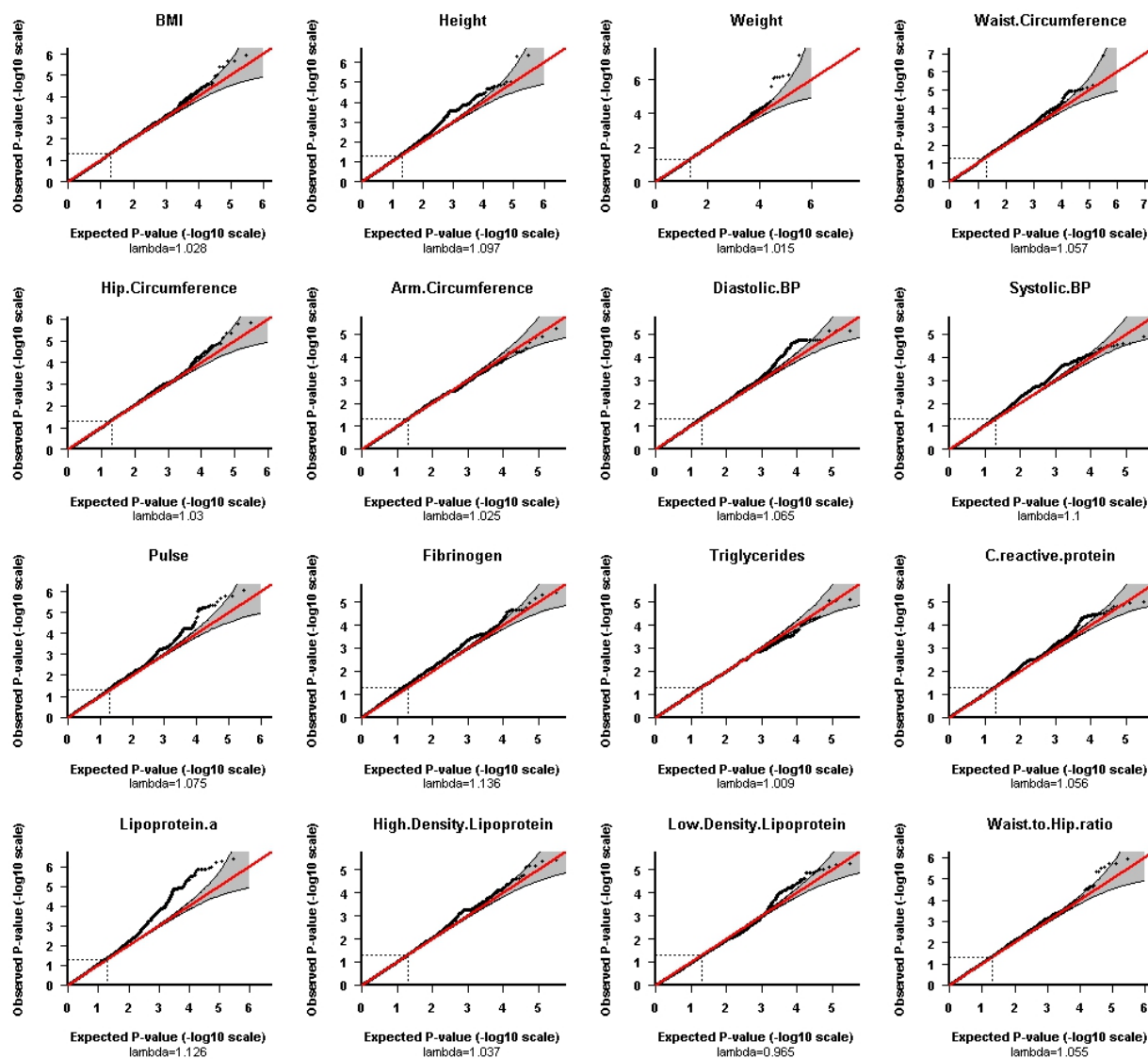
**Supplementary Figure 5:** Fitting the lengths of IBD segments in Ashkenazi Jews. Circles show the fraction of the genome in IBD segments ( $>3\text{cM}$ ) for each segment length bin (circles are at the harmonic means of bin boundaries). The green dashed line shows the best fit to a constant size history (effective size 13,600). The light blue solid line shows the best fit to a model of a bottleneck and an expansion, with the parameters shown in Supplementary Figure 4.



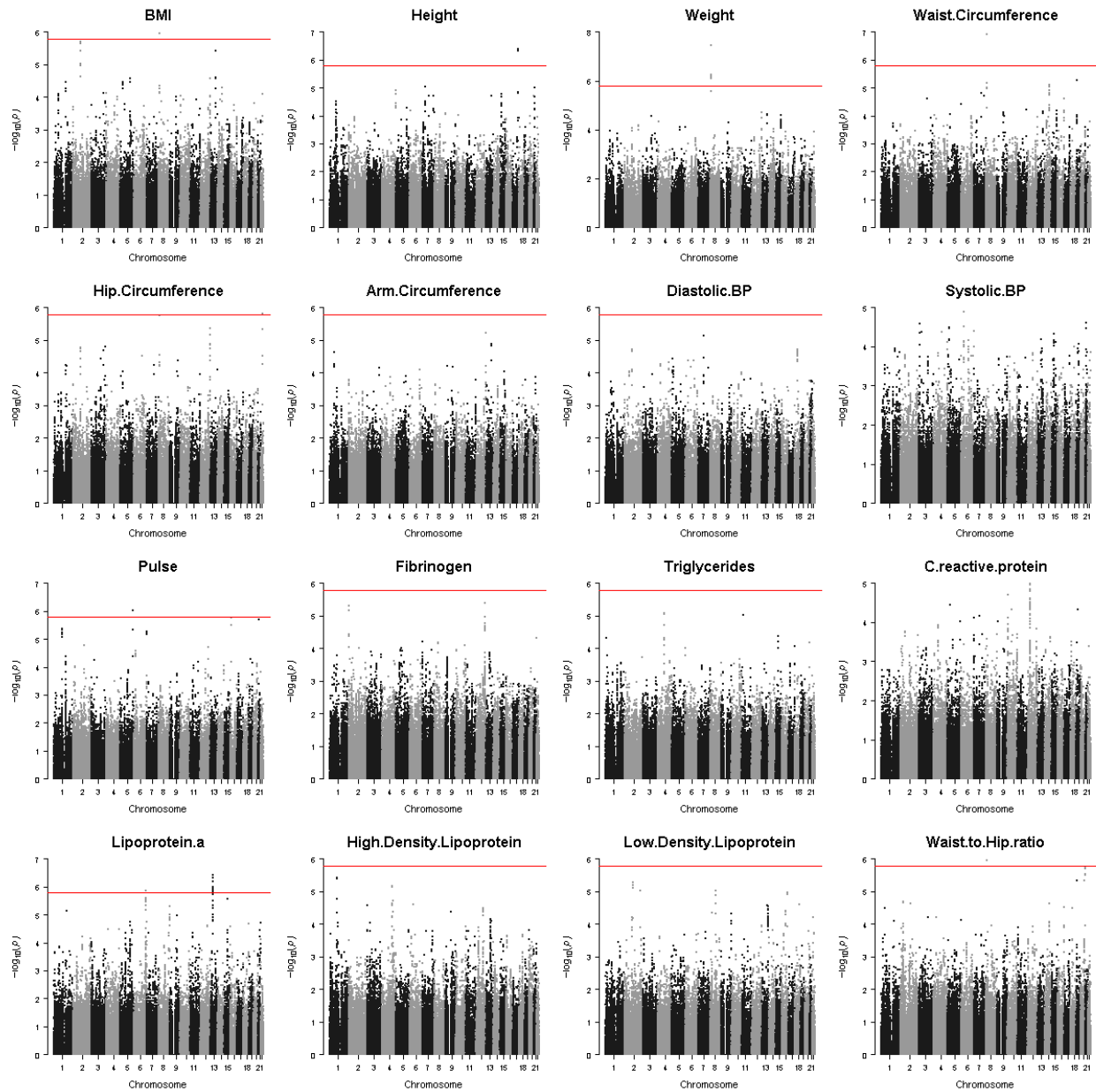
**Supplementary Figure 6:** The fitting error for each demographic parameter. The parameters in panels (A)-(D) correspond to the model shown in Supplementary Figure 4. For each parameter value, the error is the (non-weighted) sum, over all segment length bins, of the square of the log-ratio between the expected and observed fraction of the genome shared (Supplementary Figure 5). In each panel, once the designated parameter has been fixed, the error was minimized over the other three parameters.



**Supplementary Figure 7:** The number of variants with each SnpEff annotation (low, moderate, and high) across MAF ratio (KFS/Europeans) quartiles.

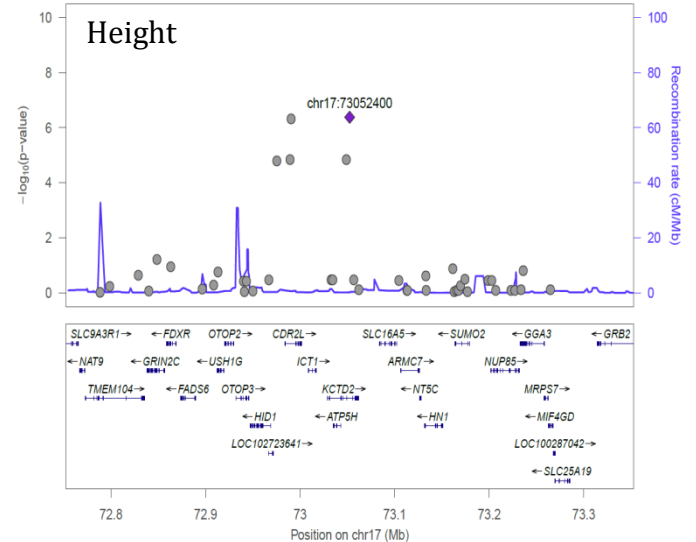
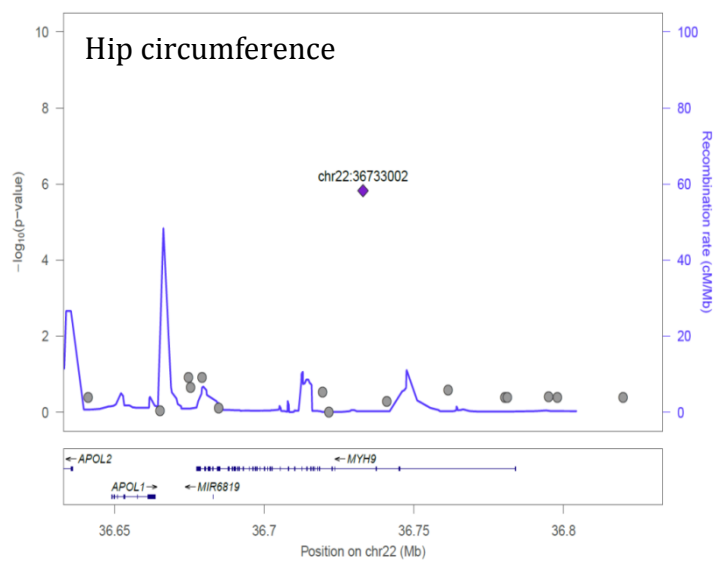
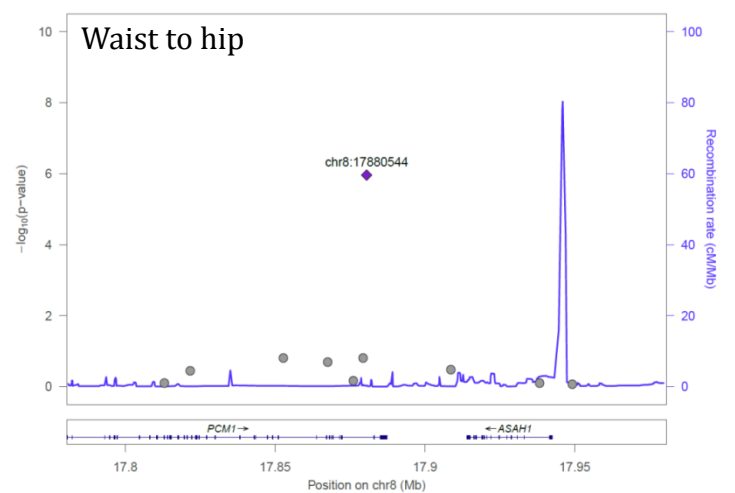
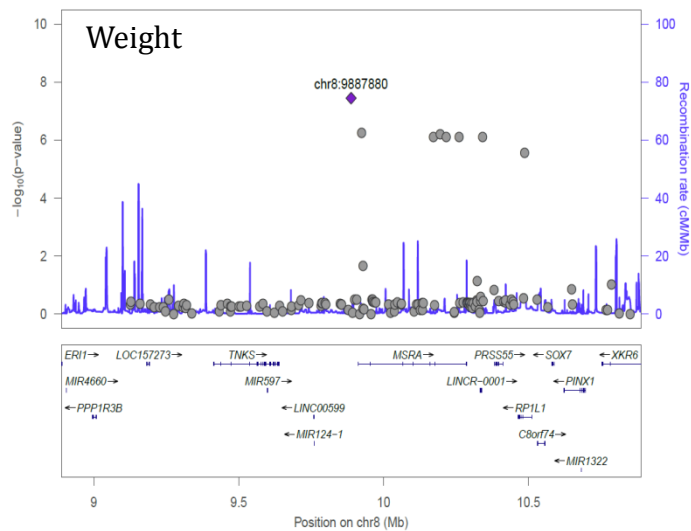
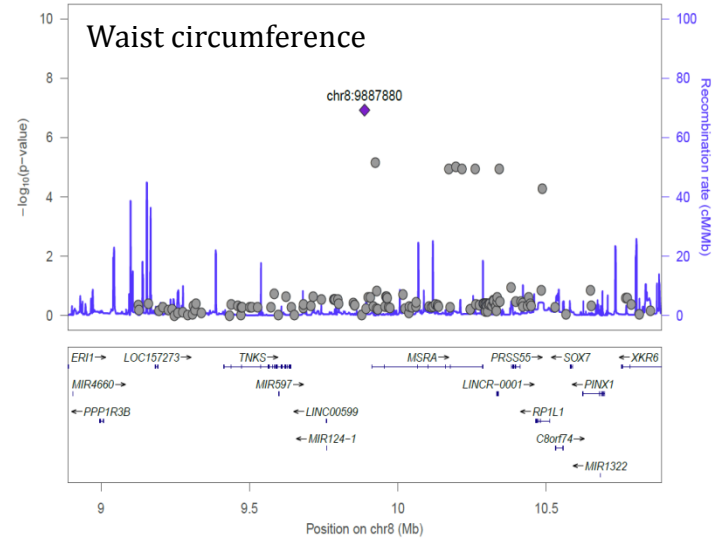
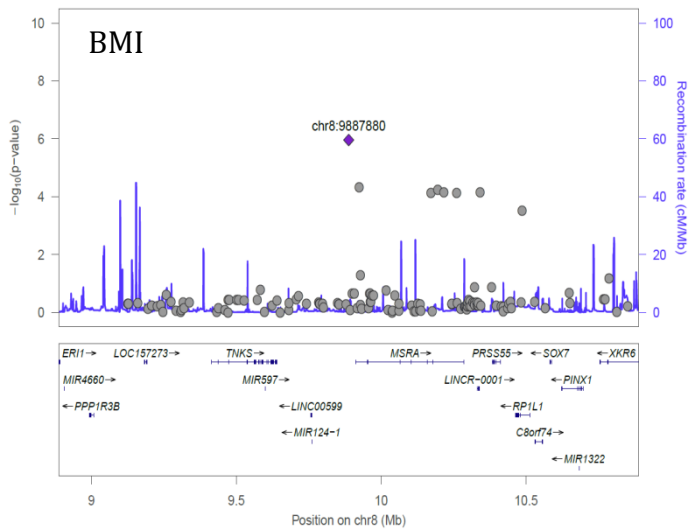


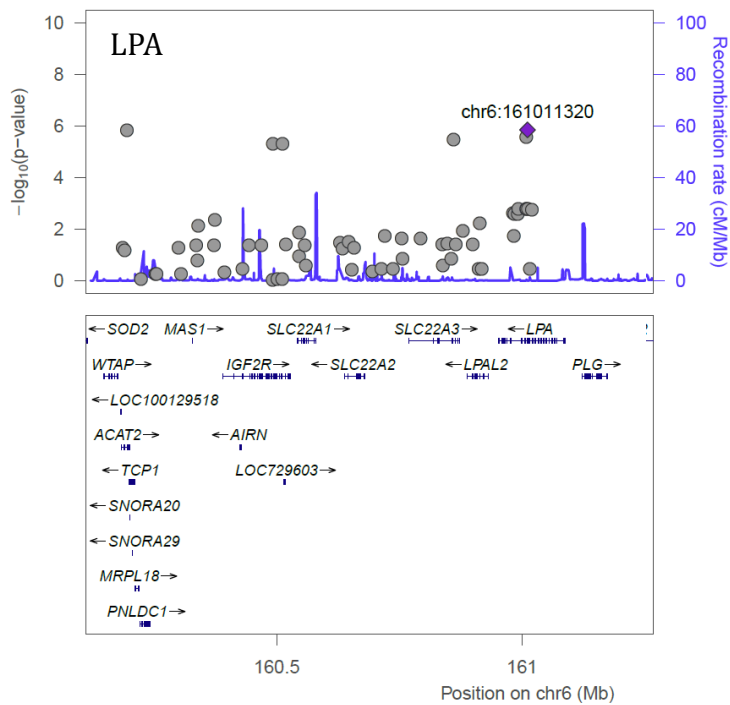
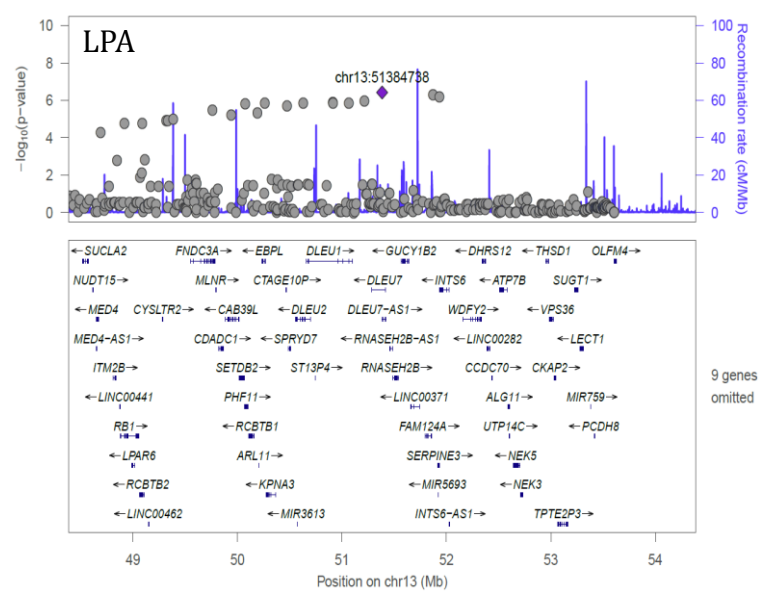
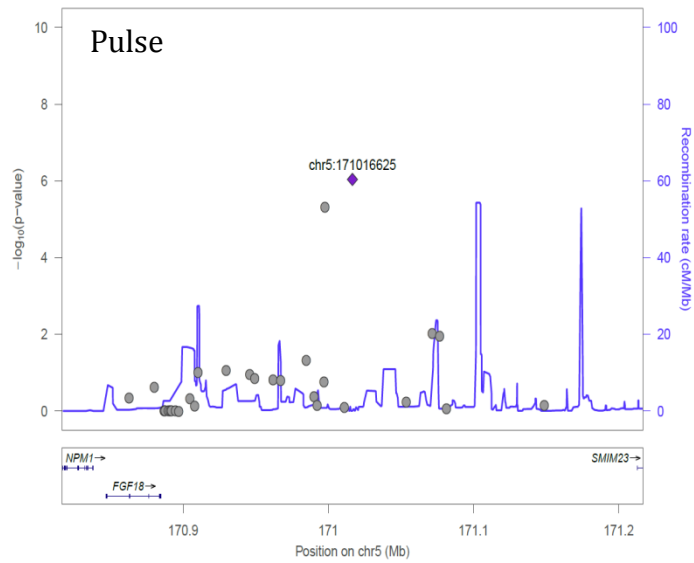
**Supplementary Figure 8:** QQ-plots of the KFS enriched variants (n=212,505) for all 16 phenotypes.



**Supplementary Figure 9:** Manhattan plots of the KFS enriched variants ( $n=212,505$ ) for all 16 phenotypes. Red line indicates a suggestive significance threshold of  $P=1.61 \cdot 10^{-6}$ .







**Supplementary Figure 10:** Locus zoom plots of top associated loci in the KFS.